

LINKING THE ETA MODEL WITH THE COMMUNITY MULTISCALE AIR QUALITY (CMAQ) MODELING SYSTEM: Ozone boundary conditions

Pius C. Lee, Jonathan E. Pleim, Rohit Mathur, Jeffery T. McQueen,
Marina Tsidulko, Geoff DiMego, Mark Iredell, Tanya L. Otte, George
Pouliot, Jeffrey O. Young, David Wong, Daiwen Kang, Mary Hart, and
Kenneth L. Schere*

INTRODUCTION

Until the recent decade, air quality forecasts have been largely based on statistical modeling techniques. There have been significant improvements and innovations made to these statistically based air quality forecast models during past years (Ryan et al., 2000). Forecast fidelity has improved considerably using these methods. Nonetheless, being non-physically-based models, the performance of these models can vary dramatically, both spatially and temporally. Recent strides in computational technology and the increasing speed of supercomputers, combined with scientific improvements in meteorological and air quality models has spurred the development of operational numerical air quality prediction models (e.g., Vaughn et al., 2004, McHenry et al., 2004).

In 2003, NOAA and the U.S. Environmental Protection Agency (EPA) signed a memorandum of agreement to work collaboratively on the development of a national air quality forecast capability. Shortly afterwards, a joint team of scientists from the two agencies developed and evaluated a prototype surface ozone concentration forecast capability for the Eastern U.S. (Davidson et al., 2004). The National Weather Service (NWS) / National Centers for Environmental Prediction (NCEP) Eta model (Black, 1994, Rogers et al., 1996, and Ferrier et al., 2003) with 12-km horizontal finite cell size was used to drive the EPA Community Multi-Scale Air Quality (CMAQ) model (Byun et al., 1999) to produce an up to 48 h Ozone (O_3) prediction. McQueen et al. (2004) and Otte et

* Pius C. Lee, Marina Tsidulko, and Mary Hart, Scientific Applications International Corporation, Camp Springs, MD. Jonathan E. Pleim, Rohit Mathur, Tanya L. Otte, George Pouliot, Jeffrey O. Young, and Kenneth L. Schere, National Oceanic and Atmospheric Administration, Research Triangle Park, NC, on assignment to the National Exposure Research Laboratory, U.S.E.P.A. Jeffery T. McQueen, Geoff DiMego, and Mark Iredell, NOAA, NWS / National Centers for Environmental Prediction, Camp Springs, MD. David Wong, Lockheed Martin Information Technology, Research Triangle Park, NC. Daiwen Kang, Science and Technology Corporation, Research Triangle Park, NC.

al. (2004) described the challenge of coupling Eta and CMAQ, and running the coupled model on a real time basis.

The general performance of the modeling system is that the system errs on over prediction (Pleim et al., 2003, Ryan et al., 2004). Figure 1 displays a typical time series of computation-domain-wide mean surface O_3 observation (AIRNOW, EPA, 2004), corresponding prediction and bias over 640 monitoring stations in Northeastern U.S.

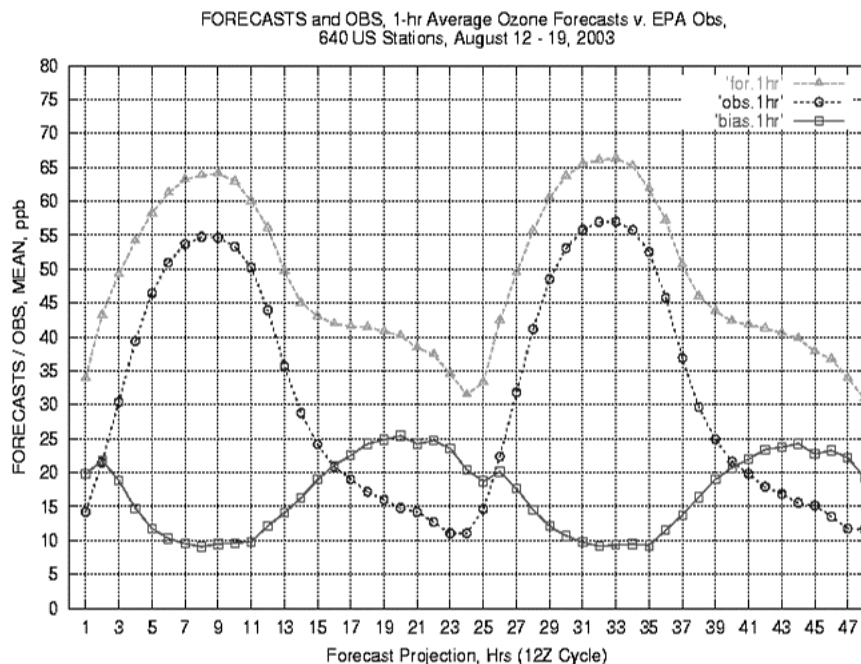


Figure 1. Mean predicted CMAQ O_3 concentration (triangle), AIRNOW observed (circle), and bias error (square) by forecast hour for all available 12 UTC cycle CMAQ predictions from August 12th to August 19th, 2003

Throughout this study, these 640 monitoring stations within the computational domain formed the basis of the performance verification (See Figure 2).

A key uncertainty in regional photochemical modeling relates to the specification of Lateral Boundary Conditions (LBC's) for O_3 and its precursor species, within both the boundary layer and the free troposphere. Specification of both temporally and spatially varying boundary conditions is desirable. While surface measurements of O_3 can be used for this purpose to provide some representation of O_3 variations along the lateral boundaries of the surface, this approach alone cannot provide information on variations within the free troposphere. An alternate approach is to use a global scale model to provide vertical variations along the lateral boundary. The current study reports on the issues and impacts associated with such an approach.

LATERAL O_3 BOUNDARY CONDITION SCHEMES

Uncertainty in the O_3 concentration at the boundaries of the computational

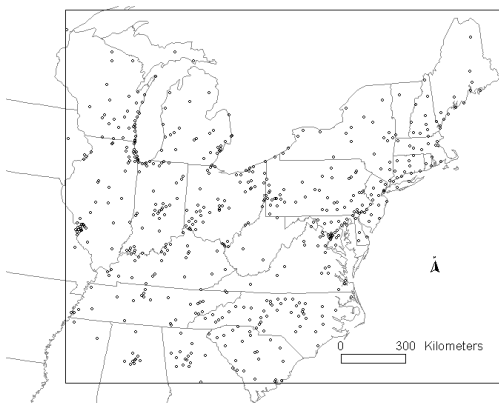


Figure 2. EPA / AIRNOW Ozone monitors in the Northeastern U.S. used in the model evaluation. Figure courtesy of the National Weather Service Meteorological Development Laboratory.

domain is one of the primary uncertainties. In the 2003 summer runs, a climatologically derived O₃ concentration profile was used. Profiles presented in Figure 3 were prescribed for each of the four boundaries and held static in time. Real-time or near real-time measurements may provide better estimates than climatologic data derived BC's. This study tested the O₃ forecasts that were generated by NCEP's spectral Global Forecast System (GFS) to derive, in part, the CMAQ O₃ LBC's. The GFS model treats O₃ as a 3-D prognostic variable (Moorthi and Iredell, 1998). It is treated as an advection trace species

with simple zonally averaged climatological derived production and depletion mechanism (Rood et al., 1991). The GFS ozone is initialized using Solar Backscatter Ultra-Violet-2 (SBUV-2) satellite observations (NCEP, NOAA 2004a). The satellite provides 12 vertical layers of O₃ concentration, with the lowest layer spanning from the surface to 250 mb. The data ingest analysis step within the GFS model system takes the O₃ field from a previous GFS forecast cycle as an initial guess and combines it with the satellite data to generate an updated O₃ field. At NCEP, both the GFS and the Eta-CMAQ model systems are run four times per day at 00, 06, 12 and 18 UTC cycles. Ideally, the GFS analysis O₃ field should be used for all CMAQ cycles. However, the GFS system starts later than the Eta-CMAQ system. There is an hour time lag for the 06 and 18 UTC cycles and a one and a half hour lag for the 00 and 12 UTC cycles between the two systems. The preparation of LBC's for CMAQ starts when Eta has finished 48 forecast hours. At that time, the GFS system is not yet ready to provide its analysis results. Hence, the GFS 6 h forecast of the previous cycle is used by the CMAQ system to derive its LBC's. The GFS system outputs O₃ every three hours on 42 sigma levels over a global 1° resolution grid (NCEP, NOAA, 2004b). The GFS O₃ field is interpolated to the CMAQ 12 km grid spatially and temporally. Another CMAQ input preparation step further extracts time varying O₃ concentration lateral BC's. Figure 4 shows an example of such GFS O₃ derived LBC's. The maxima O₃ lie in the top layer, reaching a magnitude often in excess of hundreds of ppbv.

SENSITIVITY STUDY

A base case (Case A) is defined as a run of CMAQ with its default static O₃ BC profile in Fig. 3. This CMAQ run was started at 12 UTC on May 16, 2004. A 12 UTC May 17 run was initialized using the May 16th 24 hours forecast. The target date for comparison is May 18th, 16 – 40 hours into the May 17th run.

Two sensitivity cases have been devised to investigate the impact of adopting the GFS O₃ for CMAQ's LBC specifications. First, the entire GFS O₃ column is used to derive the O₃ BC profile (Case B), replacing the default profile entirely. Figure 4 presents a sample of such profiles. It is observed that although GFS O₃ in the CMAQ top model

layer is from around 300 to around 900 ppbv,

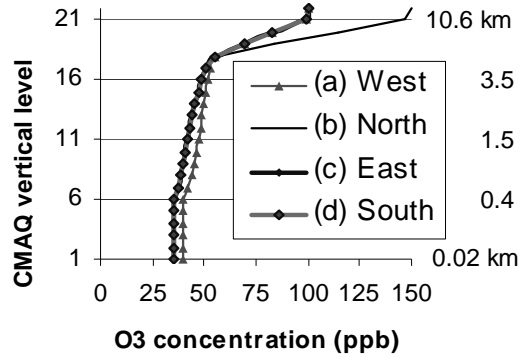


Figure 3 Ozone concentration profile used as static LBC's for Case A

in the layers below 6 km (CMAQ layer 18) the O_3 LBC profile based on GFS is usually lower in value than that based on climatologic data, except for the northern boundary. These will be time varying profiles. In Case C, the use of the GFS O_3 profile is limited to above 6 km. This reasoning stems from the fact that SBUV-2 data are not vertically resolved between the surface and 250 mb. The bulk of the total ozone measured is in the stratosphere. Therefore, the GFS has greater confidence in predicting O_3 concentrations around the tropopause and above, where there are satellite observations, than in those layers below

(NCEP, NOAA, 2004c). In the layers below, O_3 prediction depends totally on mechanical advection and a simple seasonally averaged climatologic chemical mechanism within the GFS system. The primary motivation to include O_3 in the GFS was to provide a more accurate estimate of radiative heating in the stratosphere. Therefore, the profile of O_3 below the tropopause is not an intended product of the GFS system. On the other hand, the static LBC O_3 profile of Case A is not very representative in the upper layers. The approach adopted in Case C, attempts to combine the salient features of the two data sets.

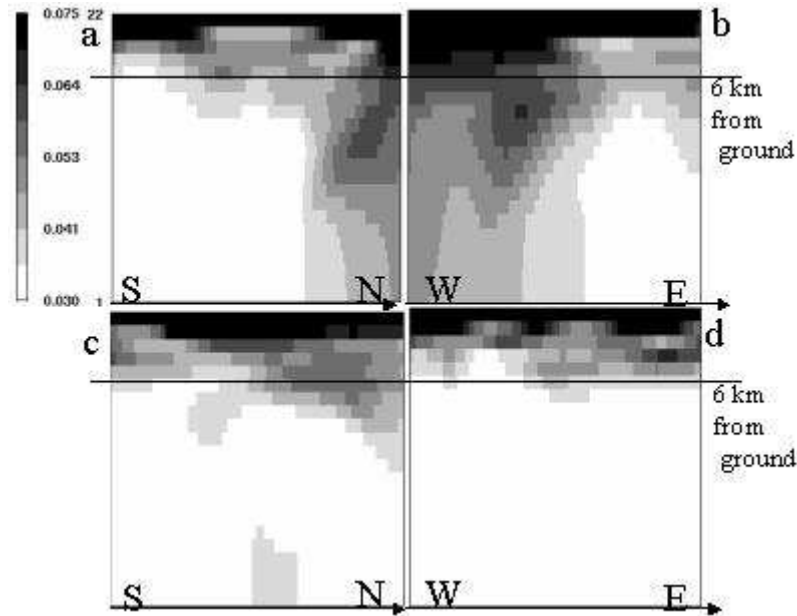


Figure 4. Samples of spatially and temporally varying GFS model derived O_3 LBC's at: (a) Western, (b) Northern, (c) Eastern, and (d) Southern boundaries respectively.

METEOROLOGICAL CONDITIONS AROUND MAY 18TH 2004

On May 16th, a fast moving cold front was migrating southeastward from the central prairies of Canada. In addition, the previous week, strong high pressure system was stationed off the U.S. Eastern Seaboard. This system hampered the eastern movement of the fast moving front. By May 18th, the front became stationary between New England and the Ohio Valley (See Figure 5a). Strong storms formed in advance of the front, and the 24 hour precipitation recorded in the southern part of Ohio registered in excess of 1 inch. In association with this frontal passage, there was rather active cumulus convection. Figure 5b shows a snapshot of the Convection Available Potential Energy (CAPE) at the surface layer as described by the initial analysis of the Eta modeling system. There was also wide-spread precipitation activity over much of the Northeastern U.S., the domain of our prototype system.

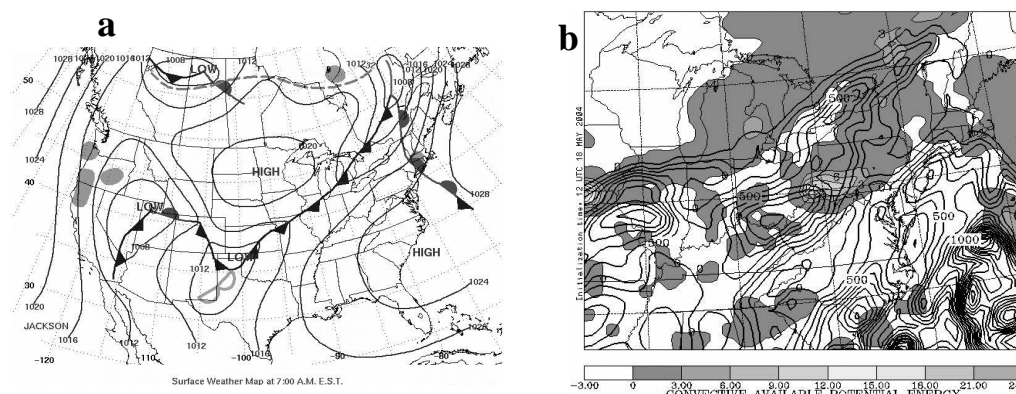


Figure 5. Meteorological conditions for May 18, 2004: (a) Surface weather map, and (b) Eta's analysis result at 6 UTC for Convective Available Potential Energy (contours at 500 J kg⁻¹ intervals) overlaid with 3 hour accumulated precipitation shaded in 3 mm interval. It indicated a maximum in southwestern PA in excess of 24 mm.

TROPOPAUSE HEIGHT PREDICTIONS

Both the GFS and the Eta models have refined vertical structure around the tropopause. Two of the objectives in obtaining a good representation of those heights are to improve forecasts of jet level winds and stratospheric-tropospheric exchange. The similarity of the two models in their description of the upper troposphere is important for the assimilation of the GFS O₃ from those layers between the two models. For instance, the activities due to deep convection would influence the sharp gradients of the O₃ concentration near the tropopause. These processes should be described in a compatible manner by the models to preserve these sharp gradients. It is thus of interest to compare the compatibility of their predictions around the tropopause. The predicted tropopause height from the GFS and Eta models on May 18, 2004 agreed well. Typically, the agreement between the two models in tropopause height prediction is rather good during most seasons. This can be partially credited to the similar vertical layer structures of the two models between 300 mb and 20 mb. The GFS model has 13 uneven layers between the heights, spaced at about 20-25 mb intervals. The Eta model has 12 layers between the same heights, spaced at about 20-29 mb intervals. Consistency in tropopause height prediction is a prerequisite to assuming that the two models are compatible in the

prediction of tropopause dynamics.

RESULTS

Figures 6 a-c present the bias in the 8-hour average maximum O_3 concentrations in the three cases forecasts valid May 18, 2004. Overall, the base Case A has the best performance (See Fig. 6a), where sporadic over predictions were clustered around the New England area and western PA. Cases B and C both showed additional clusters of over prediction just behind the cold front. Noticeably, there were additional clusters of high bias between the New England States and the Canadian border, over Lake Erie, and in western Ohio. These regions correspond to areas with prolonged, strong convective activity on that day. Between 03 and 18 UTC, the convective available potential energy at the surface in these 3 regions averaged in excess of 800, 900, and 940 $J\ Kg^{-1}$, respectively. Therefore it can be expected that the high BC O_3 concentration in the top layers derived by the GFS will be transported downwards. This is likely to have contributed to the additional high bias of Cases B and C. Namely, the frontal convective movement entrained O_3 from the model top layers to the lower layers through down drafts associated with clouds.

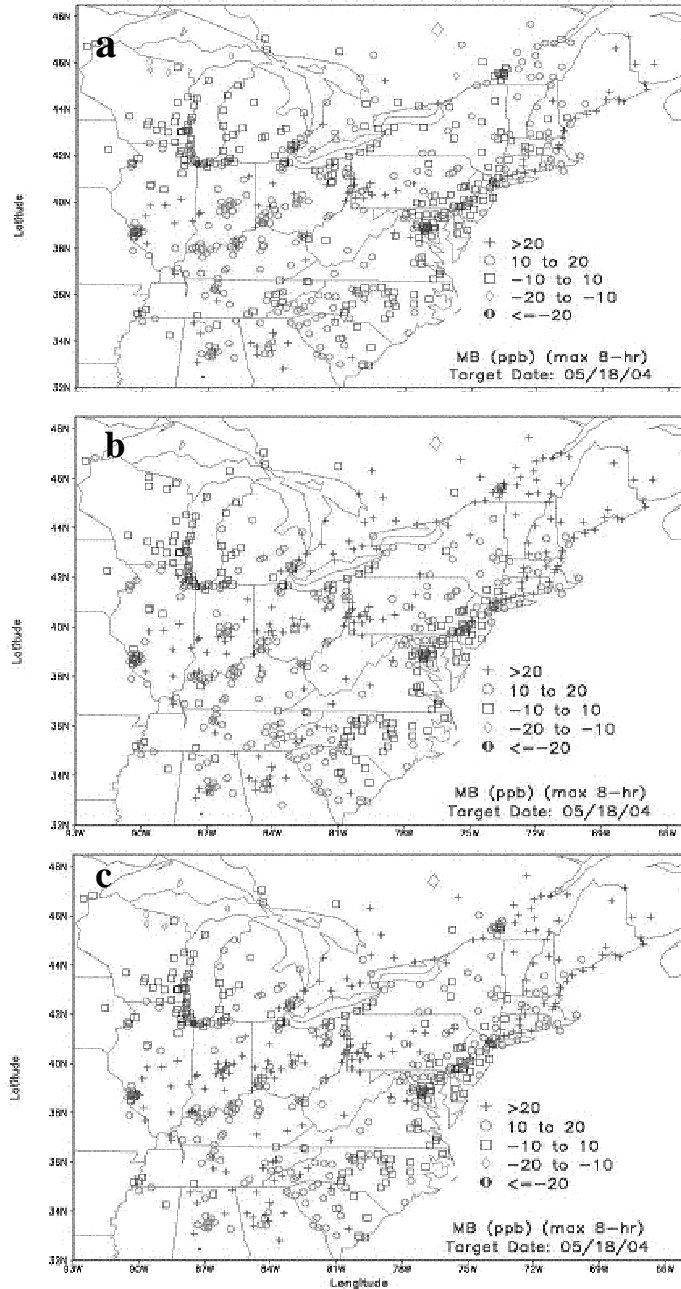


Figure 6. Distribution of 8 h average maximum O_3 mean bias for May 18, 2004 for: (a) Case A. (b) Case B. and (c) Case C.

Case C showed fewer instances of over prediction behind the cold front compared to Case B. For examples, there were fewer additional high biases just north of New York State along its border with Canada and the State of Vermont, and along the mid stretch of Lake Erie. This can be understood from the perspective that the northern boundary condition based on the GFS derived O₃ profile was often higher than that derived from climatologic average, even for the lowest layers below 6 km. Therefore, Case B tends to have the most instances of high bias in these areas close to the northern boundary. However, in the areas close to the western boundary, Case C showed more instances of high bias. In fact, Case B is the only case that had few high bias in western Illinois. This is attributed to the site's proximity to the western boundary, and to the low BC O₃ concentration provided by the GFS forecast in the layers below 6 km.

These results demonstrated three shortcomings in the use of GFS forecast based O₃ BC's. First, the vertical interpolation of the GFS O₃ profile, from the GFS's 42 to CMAQ's 22 sigma levels, is probably too coarse in the layers around the CMAQ model top. Between 300 mb and the model top of 100 mb, there are only 2 vertical layers. This insufficient resolution in CMAQ levels results in a distorted O₃ BC profile, with the peak values extending too far down into the troposphere. Second, the CMAQ vertical resolution is too coarse to describe the O₃ entrainment activity in convective areas. There are only roughly 5 vertical layers from 550 mb and the model top. Third, the dynamics of Eta may not have been adequately represented when the meteorological fields were interpolated from a 60 level step mountain vertical structure into the rather coarse sigma level structure of CMAQ. Therefore, this study suggests further investigation to explore a more consistent coupling between the various models and increased vertical resolution in CMAQ near the model top.

SUMMARY

A prototype Eta-CMAQ operational air quality forecast system had been in use throughout the summer of 2003 – Case A. Model evaluation showed that the system tended to over predict O₃. It has been proposed that the uncertainties associated with the lateral boundary condition for O₃ concentration require investigation. In these summer runs, climatologic data formed the basis for the CMAQ's LBC's, but these data are less reliable in the upper troposphere. Therefore, it was proposed that NCEP's GFS O₃ forecast be used to refine these BC's. Two schemes have been used to ingest GFS O₃ for constructing the CMAQ's LBC profiles. First, the entire GFS O₃ profile has been used to replace the profiles used in the 2003 summer runs – Case B. Second, only the part of the GFS O₃ profiles above 6 km were used – Case C. Results showed that Case A has the least amount of bias when compared to observations. It does not suggest that our confidence in using GFS O₃ profiles to derive air quality model O₃ BC profiles has been decreased. Especially in the upper troposphere, GFS has high confidence of fidelity in predicting O₃ concentrations due to good satellite observations there. It does show that a further investigation of the coupling of Eta and CMAQ in the upper layers is required.

ACKNOWLEDGEMENTS

The views expressed are those of the authors and do not necessarily represent those of the National Weather Service, NOAA or the EPA. The EPA AIRNOW program staff provided the observations necessary for quantitative model evaluation. **Disclaimer:** The research presented here was performed under the Memorandum of Understanding

between the U.S. Environmental Protection Agency (EPA) and the U.S. Department of Commerce's National Oceanic and Atmospheric Administration (NOAA) and under agreement number DW13921548. Although it has been reviewed by EPA and NOAA and approved for publication, it does not necessarily reflect their policies or views.

REFERENCES

- AIRNOW, EPA, 2004: Office of Air Quality Planning and Standards, U.S. EPA, AIRNOW Website (<http://www.epa.gov/airnow>).
- Black, T., 1994: The new NMC mesoscale Eta Model: description and forecast examples. *Wea. Forecasting*, **9**, 265-278.
- Byun, D. W., J. Young, J. Pleim, M. T. Odman, and K. Alapaty, 1999: Numerical transport algorithms for the Community Multiscale Air Quality (CMAQ) chemical transport model in generalized coordinates. Chapter 7 of *Science algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) Modeling System*. D. W. Byun and J. K. S. Ching, Eds. EPA-600/R-99/030, Office of Research and Development, U.S. Environmental Protection Agency, Washington, D.C. [Available from U.S. EPA, ORD, Washington, D.C. 20460.]
- Davidson, P. M., N. Seaman, K. Schere, R. A. Wayland, J. L. Hayes, and K. F. Carey, 2004: National air quality forecasting capability: First steps toward implementation. Preprints, *Sixth Conf. on Atmos. Chem.*, Amer. Met. Soc., Seattle, WA, 12-16 Jan 2004.
- Ferrier, B., Y. Lin, D. Parrish, M. Pondeva, E. Rogers, G. Manikin, M. Ek, M. Hart, G. DiMego, K. Mitchell, and H. Chuang, 2003: Changes to the NCEP Meso Eta Analysis and Forecast System: Modified cloud microphysics, assimilation of GOES cloud-top pressure, assimilation of NEXRAD 88D radial wind velocity data. [Available at <http://www.emc.ncep.noaa.gov/mmb/tpb.spring03/tpb.htm> or from the National Weather Service, Office of Meteorology, 1325 East-West Highway, Silver Spring, MD 20910].
- McHenry, J. N., W. F. Ryan, N. L. Seaman, C. J. Coats, Jr., J. Pudykiewicz, S. Arunachalam, and J. M. Vukovich, 2004: A real-time Eulerian photochemical model forecast system: overview and initial ozone forecast performance in the Northeast U.S. corridor. *Bull. Amer. Meteor. Soc.*, **85**, 525-548.
- McQueen, J. P., Lee, M. Tsidulko, G. DiMego, T. Otte, J. Pleim, J. Young, G. Pouliot, B. Eder, K. Schere, J. Gorline, M. Schenk, P. Dallavalle, W. Shaffer, N. Seaman, and P. Davidson, 2004: Development and Evaluation of the NOAA/EPA Prototype Air Quality Model Prediction System, Preprints, *Sixth Conf. on Atmos. Chem.*, Amer. Met. Soc., Seattle, WA, 12-16 Jan 2004.
- Moorthi, S., and Iredell, M., 1998, Prognostic Ozone: Changes to the 1998 NCEP Operational MRF Model Analysis/Forecast System: The Use of TOVS Level 1-b Radiances and Increased Vertical Diffusion. [Available at <http://www.nws.noaa.gov/om/tpb/449.htm> from the National Weather Service, Office of Meteorology, 1325 East-West Highway, Silver Spring, MD 20910].
- NCEP, NOAA, 2004a: NCEP/GFS total ozone analyses and forecasts Website (http://www.cpc.ncep.noaa.gov/products/stratosphere/strat_a_f/index.html). NOAA Air Resource Laboratory, Silver Spring, MD.
- NCEP, NOAA, 2004b: Office note 388 GRIB Website (<http://www.nco.ncep.noaa.gov/pmb/docs/on388>)
- NCEP, NOAA, 2004c: Solar Backscatter UltraViolet Instrument (SBUV/2) Website (http://www.cpc.ncep.noaa.gov/products/stratosphere/sbuv2to/sbuv2to_info.html)
- Otte, T. L., and Coauthors, 2004: Linking the Eta Model with the Community Multiscale Air Quality (CMAQ) Modeling System to build a national air quality forecasting system (submitted to *Wea. Forecasting*).
- Pleim, J., K. Schere, J. Young, G. Pouliot, T. Otte, 2003: The Models-3 Community Multi-scale Air Quality (CMAQ) Model: linking with NWS Eta Model for air quality forecasting, Proceedings, *Air Quality Focus Group Meeting*, Silver Spring, MD, 9-10 Sep 2003, pp 1-23.
- Rogers, E., T. Black, D. Deaven, G. DiMego, Q. Zhao, M. Baldwin, N. Junker, and Y. Lin, 1996: Changes to the operational "early" Eta Analysis/Forecast System at the National Centers for Environmental Prediction. *Wea. Forecasting*, **11**, 391-413.
- Ryan, W. F., C. A. Petty, and E. D. Luebehusen, 2000: Air quality forecasts in the mid-Atlantic region: current practice and benchmark skill. *Wea. Forecasting*, **15**, 46-60.
- Ryan, W. F., P. Davidson, P. Stokols, and K. Carey, 2004: Evaluation of the National Air Quality Forecasting System (NAQFS): Summary of the air quality forecasters focus group workshop. Preprints, *Sixth Conf. on Atmos. Chem.*, Amer. Met. Soc., Seattle, WA, 12-16 Jan 2004.
- Rood, R., A. R. Douglas, J. A. Kaye, M. A. Geller, C. Y. Chen, D. J. Allen, E. M. Larsen, E. R. Nash, J. E. Nielsen, 1991: Three-dimensional simulations of wintertime ozone variability in the lower stratosphere. *J. Geophys. Res.*, **96**, # D3, 5055-5071.
- Vaughn, J., and Coauthors, 2004: A numerical daily air quality forecast system for the Pacific Northwest. *Bull. Amer. Meteor. Soc.*, **85**, 549-561.